



# Linux VFAT kernel FS改良 XVFAT for kernel 2.4.20

---

2005.03.25

[machida@sm.sony.co.jp](mailto:machida@sm.sony.co.jp)

# 改良項目

---

- mount中のメディア拔去対応
  - メディア拔去のアプリケーションへの通知
  - ブロックデバイスIOエレベータキャンセル
- 日本語ファイル名サポート
- Dirty Flag対応
- TIME ZONE指定

# mount中メディア抜去対応

---

- メディアの抜去によってI/O中のファイルおよびディレクトリエントリ以外は破壊されないようにする
  - メタデータの更新を細かく管理する
- メディアの抜去後キャッシュ等無効化
  - XVFATのキャッシュはすべて無効化
  - 物理デバイスへの書き戻されていないバッファは破棄
- ファイルへのシステムコールはエラーにする
  - 抜去によるエラーなのかを検査するインターフェイスをつける   メディア交換のアプリケーションへの通知
  - ユーザランドから再マウント
- 現在は、USB MSコントローラ前提

## mount中メディア抜去対応 2

---

- デバイスへの書き込みが完了するまでシステムコールはブロック
  - アプリケーションが安全にcloseできたことを確認できる
- メディアの抜去後もXVFATを含むシステム全体が正常に動作すること
  - やって見ないとわからない...
  - いまのところ大きな問題はなし

# メディア抜去通知

---

## ○ ファイルシステムのioctl追加

- mountされてから抜かれたか検査可能
- XVFAT上のファイルのファイルディスクリプタに対して発行可能

```
ioctl(fd, VFAT_IOCTL_GETMINFO, vfat_media_info*);
```

```
#define VFAT_MINFO_CFLAG_CHANGE 1
```

```
#define VFAT_MINFO_CFLAG_INSERT 2
```

```
struct vfat_media_info{
```

```
    int change;        /* 0: not ejected, 1: ejected (read only) */
```

```
    int insert;        /* 0: not inserted, 1: inserted (read only) */
```

```
    int capability;    /* capability flags (read only) */
```

```
};
```

# メディア抜去通知 2

---

- procファイルシステムによるメディア抜去通知
  - メディアの状態
    - /proc/removabledisk/status
    - 常にメディア状態が読める
  - メディア状態の遷移
    - /proc/removabledisk/event
    - 状態が変化するとブロックが解除される

```
$ cat /proc/removabledisk/status  
sd(8,1) 1 0  
$ cat /proc/removabledisk/event  
sd(8,1) 1 1
```

# 日本語ファイル名対応

---

- JISとUnicodeで1対1に対応していない
  - 多対1になっている文字がある
  - JIS    Unicode    JISでコードが戻ってこないことがある
  - dentry cacheの動作に影響
- 上記文字コード使用時動作に問題がないことの確認
  - 必要があれば制限などをおく
  - 英字の大文字小文字同一視についても確認
- コード変換を事前に知る方法の提供
- Windowsでの変換と互換変換テーブルを利用

# 日本語ファイル名対応

---

- VFATのファイル名の文字コード変換API
  - I/Oで利用しているエンコードからVFATエントリ上の文字コード(Unicode)への変換APIを追加する
  - ディレクトリをスキャンするなどのアプリケーションでの利用を想定
  - VFAT上のopenされたディレクトリに対するディスクリプタへのioctlとする



# I/Oエレベータキャンセル

---

- 物理デバイスへのI/Oエレベータをキャンセル
  - 別々のI/Oをまとめて効率化するためにある
  - ブロックデバイスには共通に使われている
  - 書き込み順番をVFATから精密に制御するため
- 関数は取り替えられる構造になっている
  - ブロックデバイス毎に関数を定義できる
  - 現状は全部のデバイスで共有している
- VFATから関数を切り替える
  - ブロックデバイスのカーネル構造体から関数を検索し変更する
  - mount時に切り替え, umountしたら元に戻す

# dirty flag対応

---

- FAT16,32にはdirty flagがある
  - Microsoftの公式文書あり
  - mtoolsのソースに記述あり
- mount時にセット, umount時に解除する
  - XVFATとしてはこれ以上のことはしない
  - fsckやchkdsk (Windows)での利用を想定

# dirty flag対応 2

---

- FATの特殊エントリ
  - 最初の2つのエントリは予約エントリ
    - BPB\_Media bytes(0xF8)
    - End Of Clusterの値
  - 2つめにClnShutBitMaskビット
    - 1でClean, 0でdirty
  - FAT12は空きビットが無い
- MemoryStickのFATフォーマット
  - 通常のフォーマットとは範囲が違う
    - 4~64MB品まではFAT12, 128MBはFAT16
  - MS Proではメディア内に記録されている
    - メディア自身がフォーマットする
    - 現状FAT16(256M~1G)

# TIME ZONE対応

---

- タイムスタンプがずれるのを補正する
  - FATは常にlocal timeで記録している
  - kernel内は常にUTC
  - ちがうタイムゾーンのメディアをmountすると時間がおかしくなる

# TIME ZONE対応2

---

- mountオプション

- 記録時のlocal timeを与えられるように機能追加

- POSIX風指定

- zoneinfo情報はカーネル内がない
  - JST,PSTとかは無視される
- 夏時間(DST)使う場合はほぼ全部指定

# TIME ZONE対応3

---

`mount -t xvfat -o timezone=JST-9`

UTCから9時間進んでいる時間帯

`mount -t xvfat -o timezone=PST8PDT1,M4.1.0/2,M10.5.0/2`

UTCから8時間遅れている時間帯

DSTは4月の最初の日曜から10月最後の日曜まで1時間進ませる

DSTは午前2時から開始

# そのほか

---

- 独立したFile systemとして実装
  - 従来のVFATも同時に使える
  - 必ずしも上位互換でない
    - mmap(2)の書き込み禁止
    - 圧縮ファイルシステムプラグイン(CVF)禁止
- ブロックデバイスドライバにも依存
  - メディアの抜去検査ができないとどうしようもない
  - 現状USB Mass Storageのみ
- デバイスの抜去
  - デバイスがいないとメディアの存在確認ができない
  - メディアが抜かれるとデバイスが無くなる実装は対応できない